*W essendorf* ●                    ● BIOTECHNOLOGY 
                                   SYSTEMS
                                   BRANCH

# RAW SEQUENCE LISTING
# ERROR REPORT
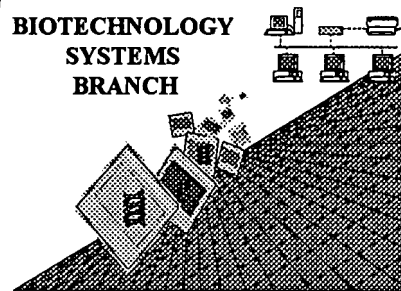
The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following CRF diskette:

Application Serial Number: ___09/096,749___

Art Unit / Team No. : ___1642___

Date Processed by STIC: ___9/9/98___

## THE ATTACHED PRINTOUT EXPLAINS THE ERRORS DETECTED.

**PLEASE BE SURE TO FORWARD THIS INFORMATION TO THE APPLICANTS BY EITHER:**

**1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANTS ALONG WITH A NOTICE TO COMPLY or,**

**2) CALLING APPLICANTS AND FAXING THEM A COPY OF THE PRINTOUT WITH A NOTICE TO COMPLY**

**THIS WILL INSURE THAT THE NEXT SUBMISSION RECEIVED FROM THEM WILL BE ERROR FREE.**

**IF YOU HAVE ANY FURTHER QUESTIONS, PLEASE CALL:**

**ARTI SHAH 703-308-4212**

# Raw Sequence Listing Error Summary

| ERROR DETECTED | SUGGESTED CORRECTION | SERIAL NUMBER: 09/096,249 |
|---|---|---|

**ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE**

1 ✓ **Wrapped Nucleics**
The number/text at the end of each line "wrapped" down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

2 _____ **Wrapped Aminos**
The amino acid number/text at the end of each line "wrapped " down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

3 _____ **Incorrect Line Length**
The rules require that a line not exceed 72 characters in length. This includes spaces.
All text must be visible on page.

4 _____ **Misaligned Amino Acid Numbering**
The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs
between the numbering. It is recommended to delete any tabs and uses spacing between the numbers.

5 _____ **Non-ASCII**
This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.

6 _____ **Variable Length**
Sequence(s) _____ contain n's or Xaa's which represented more than one residue.
As per the rules, each n or Xaa can only represent a single residue.
Please present the maximum number of each residue having variable length and
indicate in the (ix) features section that some may be missing.

7 ✓ **Wrong Designation**
Sequence(s) 1/0 contain amino acid or nucleic acid designators which are not standard
representations as per the Sequence Rules (Please refer to paragraph 1.822)

8 _____ **Skipped Sequences (OLD RULES)**
Sequence(s) _____ missing. If intentional, please use the following format for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X:
(I) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")
(xI) SEQUENCE DESCRIPTION:SEQ ID NO:X:
This sequence is intentionally skipped

Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).

9 _____ **Skipped Sequences (NEW RULES)**
Sequence(s) _____ missing. If intentional, please use the following format for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

10 _____ **Use of N's or Xaa's (NEW RULES)**
Use of N's and/or Xaa's have been detected in the Sequence Listing.
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.

11 _____ **Use of <213>Organism (NEW RULES)**
Sequence(s) _____ are missing this mandatory field or its response.

12 _____ **Use of <220>Feature (NEW RULES)**
Sequence(s) _____ are missing the <220>Feature and associated headings.
Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32)
(Sec. 1.823 of new Sequence Rules)

13 _____ **Wrong Format**
File submitted was in the alphabetical heading format of the Old Sequence Rules.This is invalid since the
"Requirements for Patent Applications Containing Nucleotide Sequence and/or Amino Acid Disclosures"
Federal Register Notice, Vol. 63, No. 104, June 1, 1998, p. 29620
applies to applications filed on or after July 1, 1998.

AKS-Biotechnology Systems Branch- 7/10/98

1642

PAGE: 1

**RAW SEQUENCE LISTING**
**PATENT APPLICATION** *US/09/096,749*

DATE: 09/09/98
TIME: 15:58:54

*INPUT SET: S28504.raw*

> This Raw Listing contains the General
> Information Section and those Sequences
> containing ERRORS.

~~Does~~ Not Comply
Corrected Diskette Needed

|    |      |                                                               |
|----|------|---------------------------------------------------------------|
| 1  |      | SEQUENCE LISTING                                              |
| 2  |      |                                                               |
| 3  | (1)  | General Information                                           |
| 4  |      |                                                               |
| 5  |      | (i) APPLICANT: Koieda, Shohei                                |
| 6  |      |                                                               |
| 7  |      | (ii) TITLE OF THE INVENTION: ARTIFICIAL ANTIBODY POLYPEPTIDES |
| 8  |      |                                                               |
| 9  |      | (iii) NUMBER OF SEQUENCES: 118                               |
| 10 |      |                                                               |
| 11 |      | (iv) CORRESPONDENCE ADDRESS:                                 |
| 12 |      | (A) ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth P.A.    |
| 13 |      | (B) STREET: 121 South Eighth Street, Ste. 1600              |
| 14 |      | (C) CITY: Minneapolis                                        |
| 15 |      | (D) STATE: MN                                                |
| 16 |      | (E) COUNTRY: USA                                             |
| 17 |      | (F) ZIP: 55402                                               |
| 18 |      |                                                               |
| 19 |      | (v) COMPUTER READABLE FORM:                                  |
| 20 |      | (A) MEDIUM TYPE: Diskette                                    |
| 21 |      | (B) COMPUTER: IBM Compatible                                 |
| 22 |      | (C) OPERATING SYSTEM: DOS                                    |
| 23 |      | (D) SOFTWARE: FastSEQ Version 2.0b                           |
| 24 |      |                                                               |
| 25 |      | (vi) CURRENT APPLICATION DATA:                               |
| 26 |      | (A) APPLICATION NUMBER: 09/096,749                           |
| 27 |      | (B) FILING DATE: June 12, 1998                               |
| 28 |      |                                                               |
| 29 |      | (vii) PRIOR APPLICATION DATA:                                |
| 30 |      | (A) APPLICATION NUMBER:                                       |
| 31 |      | (B) FILING DATE:                                             |
| 32 |      |                                                               |
| 33 |      | (viii) ATTORNEY/AGENT INFORMATION:                          |
| 34 |      | (A) NAME: Ann S. Viksnins                                    |
| 35 |      | (B) REGISTRATION NUMBER: 37,748                             |
| 36 |      | (C) REFERENCE/DOCKET NUMBER: 109.034US1                     |
| 37 |      |                                                               |
| 38 |      | (ix) TELECOMMUNICATION INFORMATION:                         |
| 39 |      | (A) TELEPHONE: (612) 373-6900                               |
| 40 |      | (B) TELEFAX: (612) 339-3061                                 |
| 41 |      |                                                               |
| 42 |      |                                                               |

**ERRORED SEQUENCES FOLLOW:**

```
      273              (2) INFORMATION FOR SEQ ID NO:13:
      274
      275          (i) SEQUENCE CHARACTERISTICS:
-->   276            (A) LENGTH: 59 base pairs
      277            (B) TYPE: nucleic acid
      278            (C) STRANDEDNESS: single
      279            (D) TOPOLOGY: linear
      280
      281          (ii) MOLECULE TYPE: cDNA
      282          (iii) HYPOTHETICAL: NO
      283          (iv) ANTI-SENSE: NO
      284          (v) FRAGMENT TYPE:
      285          (vi) ORIGINAL SOURCE:
      286
      287          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:
      288
      289    CGGGATCCCA TATGCAGGTT TCTGATGTTC CGCGTGACCT GGAAGTTGTT
-->   290    GCTGCGACC        59
      291
```

*format erro - see item 1 on Erra summary sheet*

*59*

```
      292              (2) INFORMATION FOR SEQ ID NO:14:
      293
      294          (i) SEQUENCE CHARACTERISTICS:
-->   295            (A) LENGTH: 55 base pairs
      296            (B) TYPE: nucleic acid
      297            (C) STRANDEDNESS: single
      298            (D) TOPOLOGY: linear
      299
      300          (ii) MOLECULE TYPE: cDNA
      301          (iii) HYPOTHETICAL: NO
      302          (iv) ANTI-SENSE: NO
      303          (v) FRAGMENT TYPE:
      304          (vi) ORIGINAL SOURCE:
      305
      306          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:
      307
      308    TAACTGCAGG AGCATCCCAG CTGATCAGCA GGCTAGTCGG GGTCGCAGCA
-->   309    ACAAC            55
      310
```

*same erro*

```
      311              (2) INFORMATION FOR SEQ ID NO:15:
      312
      313          (i) SEQUENCE CHARACTERISTICS:
-->   314            (A) LENGTH: 51 base pairs
      315            (B) TYPE: nucleic acid
      316            (C) STRANDEDNESS: single
      317            (D) TOPOLOGY: linear
      318
      319          (ii) MOLECULE TYPE: cDNA
      320          (iii) HYPOTHETICAL: NO
      321          (iv) ANTI-SENSE: NO
      322          (v) FRAGMENT TYPE:
      323          (vi) ORIGINAL SOURCE:
```

```
324
325            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:
326
327     CTCCTGCAGT TACCGTGCGT TATTACCGTA TCACGTACGG TGAAACCGGT G
328     51
329
```

```
402                  (2) INFORMATION FOR SEQ ID NO:20:
403
404            (i) SEQUENCE CHARACTERISTICS:
-->    405            (A) LENGTH: 55 base pairs
406            (B) TYPE: nucleic acid
407            (C) STRANDEDNESS: single
408            (D) TOPOLOGY: linear
409
410            (ii) MOLECULE TYPE: cDNA
411            (iii) HYPOTHETICAL: NO
412            (iv) ANTI-SENSE: NO
413            (v) FRAGMENT TYPE:
414            (vi) ORIGINAL SOURCE:
415
416            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:
417
418     CGGGATCCGA GCTCGCTGGG CTGTCACCAC GGCCAGTAAC AGCGTATACA
-->    419     GTGAT          55
420
```

```
529                  (2) INFORMATION FOR SEQ ID NO:27:
530
531            (i) SEQUENCE CHARACTERISTICS:
-->    532            (A) LENGTH: 51 base pairs
533            (B) TYPE: nucleic acid
534            (C) STRANDEDNESS: single
535            (D) TOPOLOGY: linear
536
537            (ii) MOLECULE TYPE: cDNA
538            (iii) HYPOTHETICAL: NO
539            (iv) ANTI-SENSE: NO
540            (v) FRAGMENT TYPE:
541            (vi) ORIGINAL SOURCE:
542
543            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:
544
545     GATCAGCTGG GATGCTCCTN NKNNKNNKNN KNNKTATTAC CGTATCACGT A
546        51
547
```

```
548                  (2) INFORMATION FOR SEQ ID NO:28:
549
550            (i) SEQUENCE CHARACTERISTICS:
-->    551            (A) LENGTH: 57 base pairs
552            (B) TYPE: nucleic acid
553            (C) STRANDEDNESS: single
```

```
554              (D) TOPOLOGY: linear
555
556          (ii) MOLECULE TYPE: cDNA
557          (iii) HYPOTHETICAL: NO
558          (iv) ANTI-SENSE: NO
559          (v) FRAGMENT TYPE:
560          (vi) ORIGINAL SOURCE:
561
562          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:
563
564     TGTATACGCT GTTACTGGCN NKNNKNNKNN KNNKNNKNNK TCCAAGCCAA
-->  565     TCTCGAT         57
566
```

```
567              (2) INFORMATION FOR SEQ ID NO:29:
568
569          (i) SEQUENCE CHARACTERISTICS:
-->  570            (A) LENGTH: 47 base pairs
571            (B) TYPE: nucleic acid
572            (C) STRANDEDNESS: single
573            (D) TOPOLOGY: linear
574
575          (ii) MOLECULE TYPE: cDNA
576          (iii) HYPOTHETICAL: NO
577          (iv) ANTI-SENSE: NO
578          (v) FRAGMENT TYPE:
579          (vi) ORIGINAL SOURCE:
580
581          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:
582
583     CTGTATACGC TGTTACTGGC NNKNNKNNKN NKCCAGCGAG CTCCAAG
584     47
585
```

```
586              (2) INFORMATION FOR SEQ ID NO:30:
587
588          (i) SEQUENCE CHARACTERISTICS:
-->  589            (A) LENGTH: 51 base pairs
590            (B) TYPE: nucleic acid
591            (C) STRANDEDNESS: single
592            (D) TOPOLOGY: linear
593
594          (ii) MOLECULE TYPE: cDNA
595          (iii) HYPOTHETICAL: NO
596          (iv) ANTI-SENSE: NO
597          (v) FRAGMENT TYPE:
598          (vi) ORIGINAL SOURCE:
599
600          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:
601
602     CATCACTGTA TACGCTGTTA CTNNKNNKNN KNNKNNKTCC AAGCCAATCT C
603        51
604
```

*INPUT SET: S28504.raw*

```
2099                    (2) INFORMATION FOR SEQ ID NO:110:
2100
2101            (i) SEQUENCE CHARACTERISTICS:
-->  2102            (A) LENGTH: 94 amino acids
2103                (B) TYPE: amino acid
2104                (C) STRANDEDNESS: single
2105                (D) TOPOLOGY: linear
2106
2107            (ii) MOLECULE TYPE: peptide
2108            (iii) HYPOTHETICAL: NO
2109            (iv) ANTI-SENSE: NO
2110            (v) FRAGMENT TYPE: internal
2111            (vi) ORIGINAL SOURCE:
2112
2113            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:
2114
2115
2116   Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr Ser Leu Leu
2117    1           5                  10              15
2118   Ile Ser Trp Asp Ala Pro Ala Val Thr Val Arg Tyr Tyr Arg Ile Thr Tyr Gly Glu
2119   20              25              30              35
-->  2120  Thr Gly Gly Asn Ser Pro Val Oxx Glu Phe Thr Val Pro Gly Ser Lys Ser Thr Ala
2121    40              45              50              55
2122   Thr Ile Ser Gly Leu Lys Pro Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr
2123        60              65              70              75
2124   Gly Arg Gly Asp Ser Pro Ala Ser Ser Lys Pro Ile Ser Ile Asn Tyr Arg Thr
2125        80              85              90
2126
```

→ *invalid amino acid - see 1.822(b) of*
*Sequence Rules*
*(see item*
*7 or Error*
*Summary*
*sheet)*

```
2127                    (2) INFORMATION FOR SEQ ID NO:111:
2128
2129            (i) SEQUENCE CHARACTERISTICS:
-->  2130            (A) LENGTH: 248 base pairs
2131                (B) TYPE: nucleic acid
2132                (C) STRANDEDNESS: single
2133                (D) TOPOLOGY: linear
2134
2135            (ii) MOLECULE TYPE: cDNA
2136            (iii) HYPOTHETICAL: NO
2137            (iv) ANTI-SENSE: NO
2138            (v) FRAGMENT TYPE:
2139            (vi) ORIGINAL SOURCE:
2140
2141            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 111:
2142
2143    CTGCTGATCA GCTGGGATGC TCCTGCAGTT ACCGTGCGTT ATTACCGTAT
2144   CACGTACGGT
2145
2146    GAAACCGGTG GTAACTCCCC GGTTCAGGAA TTCACTGTAC CTGGTTCCAA
2147   GTCTACTGCT
2148
2149    ACCATCAGCG GCCTGAAACC GGGTGTCGAC TATACCATCA CTGTATACGC
```

*60  where are*
*cumulative*
*base totals*
*at end of*
*each line ?*

```
        2150   TGTTACTGGC
        2151
        2152    CGTGGTGACA GCCCAGCGAG CTCCAAGCCA ATCTCGATTA ACTACCGTAC
        2153   CTAGTAACTC
        2154
-->     2155   GAGGATCC                                               248
```

# SEQUENCE VERIFICATION REPORT
## PATENT APPLICATION  *US/09/096,749*

*INPUT SET: S28504.raw*

| Line | Error | Original Text |
|------|-------|---------------|
| 276 | Entered (59) and Calc. Seq. Length (9) differ | (A) LENGTH: 59 base pairs |
| 290 | # of Sequences for line conflicts w/ running total | GCTGCGACC    59 |
| 295 | Entered (55) and Calc. Seq. Length (5) differ | (A) LENGTH: 55 base pairs |
| 309 | # of Sequences for line conflicts w/ running total | ACAAC        55 |
| 314 | Entered (51) and Calc. Seq. Length (0) differ | (A) LENGTH: 51 base pairs |
| 405 | Entered (55) and Calc. Seq. Length (5) differ | (A) LENGTH: 55 base pairs |
| 419 | # of Sequences for line conflicts w/ running total | GTGAT        55 |
| 532 | Entered (51) and Calc. Seq. Length (0) differ | (A) LENGTH: 51 base pairs |
| 551 | Entered (57) and Calc. Seq. Length (7) differ | (A) LENGTH: 57 base pairs |
| 565 | # of Sequences for line conflicts w/ running total | TCTCGAT      57 |
| 570 | Entered (47) and Calc. Seq. Length (0) differ | (A) LENGTH: 47 base pairs |
| 589 | Entered (51) and Calc. Seq. Length (0) differ | (A) LENGTH: 51 base pairs |
| 2102 | Entered (94) and Calc. Seq. Length (93) differ | (A) LENGTH: 94 amino acids |
| 2120 | Wrong Amino Acid Designator | Thr Gly Gly Asn Ser Pro Val Oxx Glu Phe Thr Val Pro Gly |
| 2130 | Entered (248) and Calc. Seq. Length (8) differ | (A) LENGTH: 248 base pairs |
| 2155 | # of Sequences for line conflicts w/ running total | GAGGATCC |
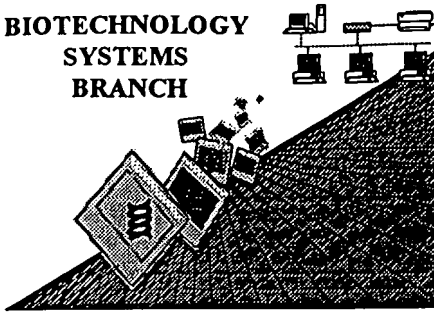
# Notice of Availability of
# Checker Program

Applicant Aid for Biotechnology Computer Readable Form (CRF)
Sequence Listing Submissions

The Patent and Trademark Office (PTO) has developed a computer program, called Checker, that will aid applicants in identifying and correcting errors prior to making submissions for compliance with the Requirements for Patent Applications Containing Nucleotide Sequence and/or Amino Acid Sequence Disclosures (Sequence Rules: 37CRF 1.821 through 1.825).
Final rules were published in the *Federal Register* (55 FR18230) on May 1, 1990, and in the PTO *Official Gazette* (1114 Off.Gaz.PatOffice 29) on May 15, 1990.

Checker is a DOS-based software program that is intended to assist users in determining whether errors may be present in the sequence listings, and is not intended to guarantee that the submission is error-free.

The most current version of the software is available via computer downloading, details are below. Copies on diskette are also available. Updated software versions will not be automatically mailed out; any updates will be announced in the PTO *Official Gazette*.

The software can be accessed/requested from the following locations:

1) Dial-up access through the Internet.  Location is ftp://ftp.uspto.gov
   The software is in current directory:  pub/checker/
   Download all the files.  Cost:  Free-of-charge

3) For diskette copies, mail to:   U.S.P.T.O., OEIP, CRYSTAL PARK 3, SUITE 441
                                    WASHINGTON DC  20231

   COST FOR DISKETTE IS  $ 25.00
   METHOD OF PAYMENT:
        Check payable to Commissioner of Patents and Trademarks
        VISA/ Mastercard/ Charge- Charges can be faxed to 703-306-2737
        PTO Deposit Account

*For Further Information, Contact:* **Arti Shah at 703-308-4212**